Nucleic Acids Research, Vol. 20, No. 6 1251–1257

The gene coding for small ribosomal subunit RNA in the basidiomycete *Ustilago maydis* contains a group I intron

Rupert De Wachter, Jean-Marc Neefs, Anne Goris and Yves Van de Peer
Departement Biochemie, Universiteit Antwerpen (UIA), Universiteitsplein 1, B-2610 Antwerpen, Belgium

Received January 6, 1992; Accepted February 15, 1992

EMBL accession no. X62396

ABSTRACT

The nucleotide sequence of the gene coding for small ribosomal subunit RNA in the basidiomycete *Ustilago maydis* was determined. It revealed the presence of a group I intron with a length of 411 nucleotides. This is the third occurrence of such an intron discovered in a small subunit rRNA gene encoded by a eukaryotic nuclear genome. The other two occurrences are in *Pneumocystis carinii*, a fungus of uncertain taxonomic status, and *Ankistrodesmus stipitatus*, a green alga. The nucleotides of the conserved core structure of 101 group I intron sequences present in different genes and genome types were aligned and their evolutionary relatedness was examined. This revealed a cluster including all group I introns hitherto found in eukaryotic nuclear genes coding for small and large subunit rRNAs. A secondary structure model was designed for the area of the *Ustilago maydis* small ribosomal subunit RNA precursor where the intron is situated. It shows that the internal guide sequence pairing with the intron boundaries fits between two helices of the small subunit rRNA, and that minimal rearrangement of base pairs suffices to achieve the definitive secondary structure of the 18S rRNA upon splicing.

INTRODUCTION

Group I introns have been hitherto observed in the nuclear genomes of eukaryotes, and in the genomes of mitochondria, plastids, a cyanelle, cyanobacteria, and the bacteriophage T4. In eukaryotes their presence has been noted much more frequently in organellar than in nuclear genomes. They have been found in species belonging to the ciliates, the green algae, the plants, the fungi, and in the slime mould *Physarum polycephalum*. A distribution of presently known group I introns over the different genome types and genes where they occur can be found in Table 1.

In several instances, group I introns have been shown to be self-splicing *in vitro*, and detailed models have been derived for the splicing mechanism and for the RNA structures involved in it (1–4). Some group I introns comprise, or overlap with, an open reading frame that codes for a maturase or for an endonuclease. The function of maturases is to facilitate *in vivo* splicing. These enzymes, however, can also be coded by extraneous genes, which is the case of certain mitochondrial group I introns where splicing is assisted by maturases derived from nuclear genes (5). Some group I introns are mobile, i.e. they are able to move into intronless copies of the same gene on the occasion of a genetic cross. Mobile introns comprise an open reading frame coding for an endonuclease, which recognizes and cuts a specific target sequence, thus allowing insertion of the intron by a mechanism probably akin to transposition (6). Possibly, the same mechanism also operates in the movement of the introns to other genes, and even to other species, accounting for the widespread occurrence of group I introns in unrelated genes and in species that are distant in evolutionary terms. Non-mobile group I introns then could arise by any mutation that removes or damages the endonuclease-coding sequence while leaving intact the structures essential for splicing. However, an alternative hypothesis (7) explains the movement of introns to new sites by a mechanism involving reverse splicing followed by reverse transcription, which does not require the action of an intron-encoded endonuclease.

As can be seen in Table 1, about one fifth of the hitherto discovered group I introns are located in large subunit (LSU) ribosomal DNA sequences of eukaryotic nuclear, mitochondrial, and plastidial genomes. As many as 6 group I introns are present in the plastid LSU rDNA of the green alga *Chlamydomonas eugametos* (8). Group I introns seem to occur much more sparingly in the DNA coding for the small subunit (SSU) ribosomal RNA. Only three cases have been reported although the number of published SSU rDNA sequences exceeds that of LSU rDNA sequences by a factor eight (9, 10).

In the eukaryotic nuclear genome, ribosomal RNA genes are the only ones hitherto found to contain group I introns. Nuclear LSU rDNA contains a single intron in certain species of the ciliate genus *Tetrahymena*, and 2 or 3 introns, depending on the strain, in the slime mould *P. polycephalum*. In nuclear SSU rDNA, the reported occurrences are in the fungus *Pneumocystis carinii* (11) and the green alga *Ankistrodesmus stipitatus* (12). The third case of an SSU rDNA containing a group I intron is in the plastidial genome of the green alga *Chlamydomonas moewusii* (13). Each of these SSU rDNAs contains a single intron, each time in a different location. In the framework of a study of fungal phylogeny based on small ribosomal subunit RNA sequencing, we discovered the presence of a group I intron in the nuclear SSU rDNA of the basidiomycetous yeast *Ustilago maydis*. The structure of the 18S rRNA and its intron is described below.

MATERIALS AND METHODS

The yeast *Ustilago maydis*, strain MUCL 30488 (Mycothèque de l'Université catholique de Louvain, Louvain-la-Neuve, Belgium; corresponding to strain CBS 445.63, Centraal Bureau voor Schimmelcultures, Baarn, the Netherlands) was cultivated in 3 l of 2% brewery malt extract at 25°C. Cells were collected by centrifugation, washed with 0.01 M sodium phosphate buffer pH 7, and kept frozen at −70°C until further use. High molecular weight DNA was isolated by grinding 5.8 g of cells in a mortar in the presence of 1 g alumina and liquid nitrogen. The crushed cells were suspended in a buffer containing 0.05 M Tris, 0.01 M $MgCl_2$, 0.05 M NaCl, 10 g/l SDS, adjusted to pH 7.4 with HCl. Phenol extraction and precipitation of nucleic acids were carried out as described by Sambrook et al. (14).

The DNA was digested with a set of restriction enzymes followed by agarose gel electrophoresis, Southern blotting, and hybridization with a probe specific for small ribosomal subunit RNA (15). Digestion with the enzyme Pst I yielded a single fragment of 8 kb containing the entire 18S rRNA gene, which was ligated into the plasmid vector pUC18. Transformation, colony hybridization, recovery of the 18S rDNA-containing plasmid, and sequence analysis of the 18S rDNA were performed as previously described (15).

An alignment of conserved sequences in group I introns as published by Michel and Westhof (4) was extended to 101 group I introns. A matrix of dissimilarities, corrected for multiple substitutions (16), was computed (17). From this matrix a dendrogram was constructed according to the neighbour joining method (18).

RESULTS

The sequence determined for *U. maydis* SSU rDNA has been submitted to the EMBL nucleotide sequence library and given the accession number X62396. It was fitted into an alignment of small ribosomal subunit RNA sequences that we maintain and update regularly (9). At the moment this alignment contains data on the primary and secondary structure of cytoplasmic SSU rRNAs of 51 fungi (19, 20; unpublished results), not including oomycetes and slime moulds. This set comprises 37 ascomycetes, 8 basidiomycetes, 4 chytridiomycetes, 1 zygomycete, and a fungus of uncertain taxonomic status, *P. carinii*. The chain length of these fungal 18S rRNAs ranges from 1759 to 1842 nucleotides, the length variation being due to the presence or absence of short sequences in a restricted number of areas of variable primary and secondary structure. In the *U. maydis* SSU rDNA, on the contrary, the nucleotides corresponding to the 5′- and 3′-termini of the 18S rRNA embrace a sequence of 2209 nucleotides. Alignment showed the extra length to be due to a 411 nucleotide insertion between the nucleotides occupying positions 1158 and 1159 counting from the 5′-terminus of the SSU rRNA. This situation is reminiscent of those described for the plastid SSU rDNA of the green alga *C. moewusii* (13), and for the nuclear SSU rDNAs of the fungus *P. carinii* (11) and the green alga *A. stipitatus* (12). In these cases, inserts of respectively 402, 390, and 394 nucleotides, each occurring in a different area of the SSU rDNA, were identified as group I introns. The *U. maydis* SSU rDNA insert sequence was therefore examined for the presence of consensus sequences and for the existence of a potential secondary structure characteristic for group I introns (3, 21).

The complete nucleotide sequence determined for the 18S rRNA of *U. maydis*, excluding the 411 nucleotide insert, is shown in Fig. 1 in the form of the secondary structure model described by Neefs et al. (9). This model is identical to that derived earlier by Gutell et al. (22), except for the folding in variable area V3 and the fact that a specific secondary structure is proposed for variable area V4 (23). The 411 nucleotide insert, which has no equivalent in other SSU rRNA sequences, is situated between helices 31 and 32, in an area of conserved primary and secondary structure. The secondary structure model derived for this insert is shown in Fig. 2a. The latter model identifies the insert as a group I intron since it possesses all its characteristic features (21), viz. the conserved sequences P, Q, R, S, the helices P1 to P9, and the pseudoknot structure formed by helices P3 and P7. Helix P9.0, demonstrated by Michel et al. (3) is present as well. The insert is also capable of forming a structure, shown in Fig. 2b and known as helix P10, where a section of the 3′-strand of hairpin P1 engages in alternative pairing with the first nucleotides following the intron-exon boundary, thus providing an internal guide sequence that brings the nucleotides bordering the splice site in close vicinity (1). The intron does not contain an open reading frame of significant length, a property it shares with other group I introns discovered in the nuclear rRNA genes, with the exception of the LSU rDNA intron 3 in *P. polycephalum* (24).

DISCUSSION

Presence of group I introns in nuclear rRNA genes

The structure of the 411 nucleotide insert in *U. maydis* SSU rDNA identifies it as a group I intron, even though this has not yet been proven experimentally by examination of the 18S rRNA structure or by in vitro splicing experiments of an rRNA precursor. The presence of group I introns in nuclear SSU rDNA had previously been discovered in just two species, the fungus

Table 1. Distribution of known group I introns: numbers found in different genes and genome types (a).

Gene product (b)	Genome type						
	Nuclear	Mitoch.	Plastid	Cyanelle	Cyanob.	Phage	Total
LSU rRNA	4 (c)	7	12				23
SSU rRNA	2		1				3
tRNA			10	1	5		16
Photosystem II			6				6
Cytochrome c oxidase		28					28
Cytochrome b		11					11
NADH dehydrogenase subunits		15					15
ATPase subunits		3					3
Miscellaneous						4	4
Total		64	29	1	5	4	109 (d)

(a)Most of the papers reporting intron structures are cited in references 4 and 6. More recent reports can be found in references 12, 27, 30, and 31.
(b)Abbreviations used: LSU rRNA, large ribosomal subunit RNA; SSU rRNA, small ribosomal subunit RNA.
(c)The intron present at the same site in the LSU rDNA of 6 different *Tetrahymena* species (32) is counted as one item.
(d)The alignment of conserved group I intron sequences which served to construct the tree of Fig. 3 was limited to 101 introns because sequences were not available for a number of introns in the following genes: *Chlamydomonas moewusii* chloroplast LSU rDNA (cited in reference 6), *Chlamydomonas reinhardtii* chloroplast *psb*A (33), and *Schizosaccharomyces pombe* mitochondrion *cox*I (34). The intron of *Scenedesmus obliquus* mitochondrion LSU rDNA (27) was also omitted because no secondary structure model was available.

P. carinii (11) and the green alga *A. stipitatus* (12). In nuclear LSU rDNA, they have been found in certain species of the ciliate genus *Tetrahymena*, and in the slime mould *P. polycephalum*. The occurrence of group I introns in widely divergent species is thought to be due to lateral gene transfer (5, 6, 25). Transposition at the DNA level, as described by Dujon (6), requires an endonuclease coded by the mobile intron itself. The fact that more than half of the presently known group I introns

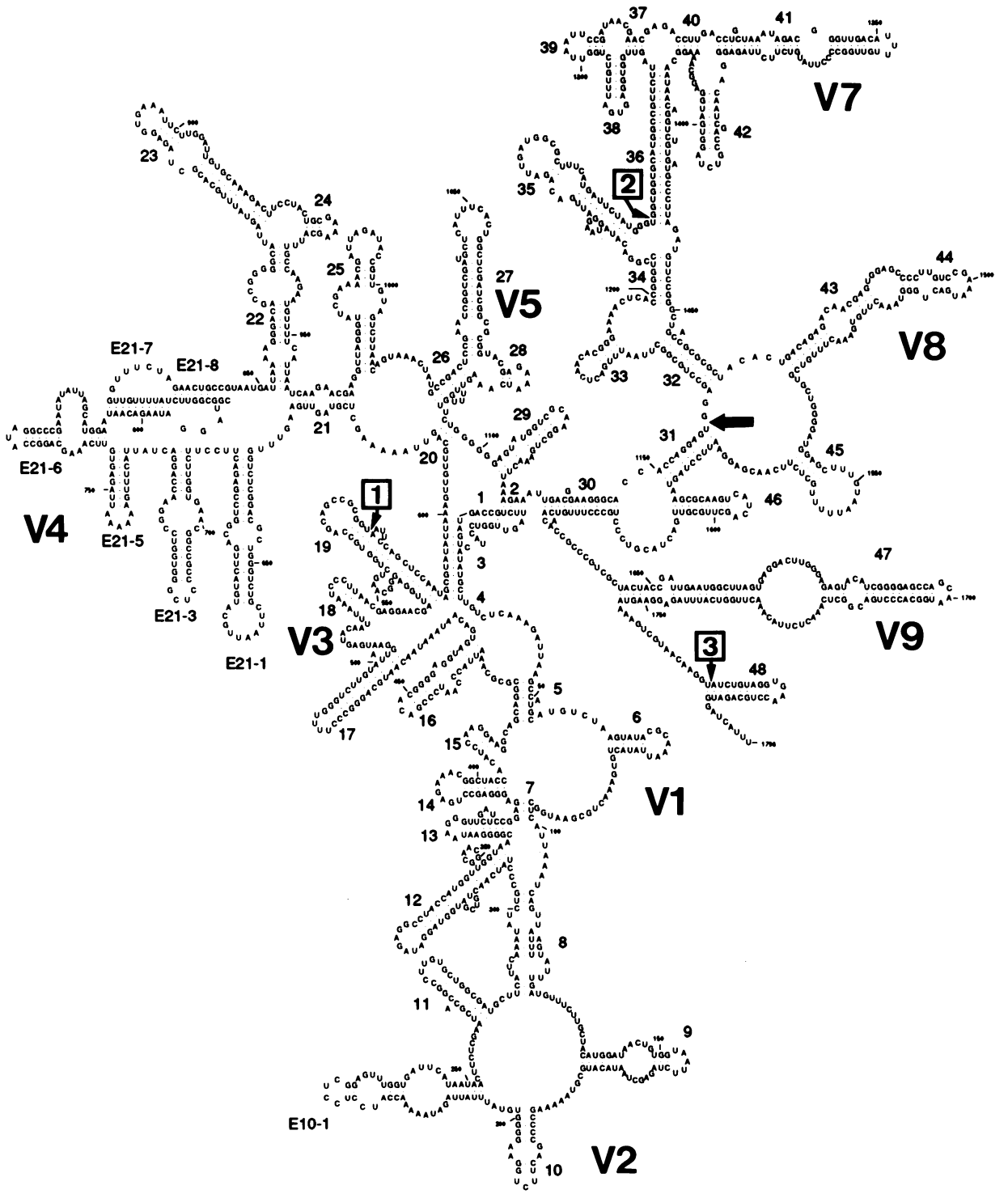

Fig. 1. Nucleotide sequence and secondary structure model for *U. maydis* SSU rRNA. Helix numbering is according to reference 9. V1 to V9 indicate areas of variable primary and secondary structure. The insertion site of the 411 nucleotide group I intron is indicated by a bold arrow. Numbered arrows indicate the sites homologous to the group I intron insertion sites in the plastidial SSU rRNA of *C. moewusii* (arrow 1) and in the nuclear-encoded SSU rRNAs of the green alga *A. stipitatus* (arrow 2) and the fungus *P. carinii* (arrow 3).

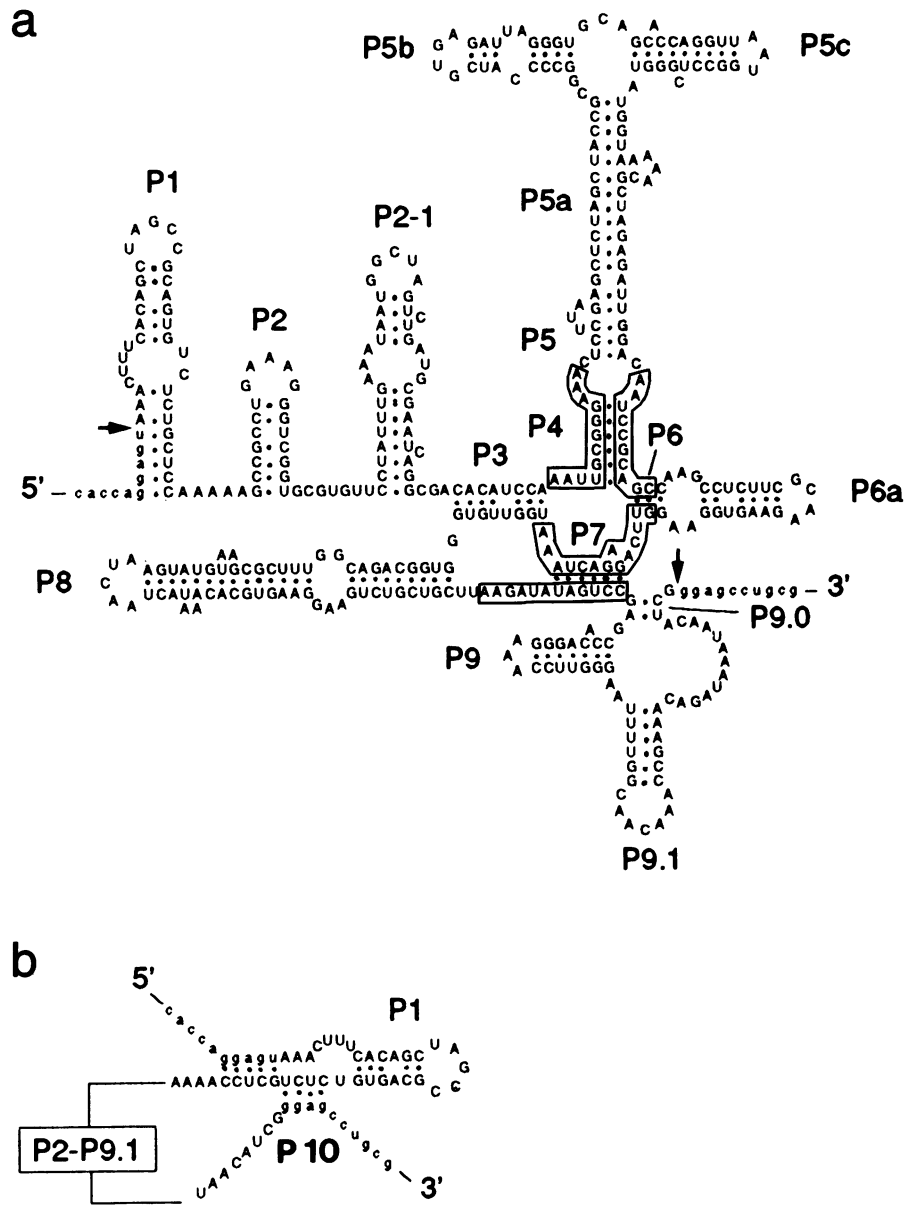

Fig. 2. Structure of the *U. maydis* SSU rRNA group I intron. a) Nucleotide sequence and secondary structure model, drawn according to the conventions recommended by Burke et al. (21), except for helices P9.n, which are drawn such as to show the presence of the helix numbered P9.0 by Michel et al. (3). The boxed sequences are, from 5' to 3', conserved sequence elements P, Q, R, S. Nucleotides to the left and to the right of the intron and belonging to the mature SSU rRNA are written in lower case. Splice sites are indicated by arrows. b) Structure of helix P10 comprising the internal guide sequence.

in rRNA genes do not possess an open reading frame is not in conflict with this hypothesis, since the endonuclease-coding gene may have been deleted later in evolution, resulting in loss of mobility but not of splicing ability. However, the presence of these introns is more intriguing in nuclear genes than in bacterial or organelle genes, since the expression of the endonuclease gene is much easier to envision if the transposon is in direct contact with a protein-synthesising system. This conceptual difficulty does not exist with the hypothesis of Woodson and Cech (7), which explains transposition to a new site by reverse splicing occurring at the RNA level.

Relatedness to other group I intron sequences

The *U. maydis* SSU rDNA intron can be classified as belonging to subdivision IB as defined by Michel et al. (26) on the basis of the absence of additional hairpins between helices P3 and P7, and a number of other characteristics enumerated by Cech (2). More recently, Michel and Westhof (4) published an alignment of 87 group I intron sequences. They computed distances on the basis of nucleotide changes at 83 alignment positions corresponding to the conserved core of the intron structure. On this basis they derived a more elaborate subdivision of the introns into groups IA to ID, some of which are further divided into subgroups.

We extended this alignment covering 83 positions to include newly published group I intron sequences for which a secondary structure model was available. An evolutionary relatedness tree was constructed from the resulting alignment of 101 sequences as described in the methods section. The complete tree (not shown) reveals a number of clusters, some of which coincide

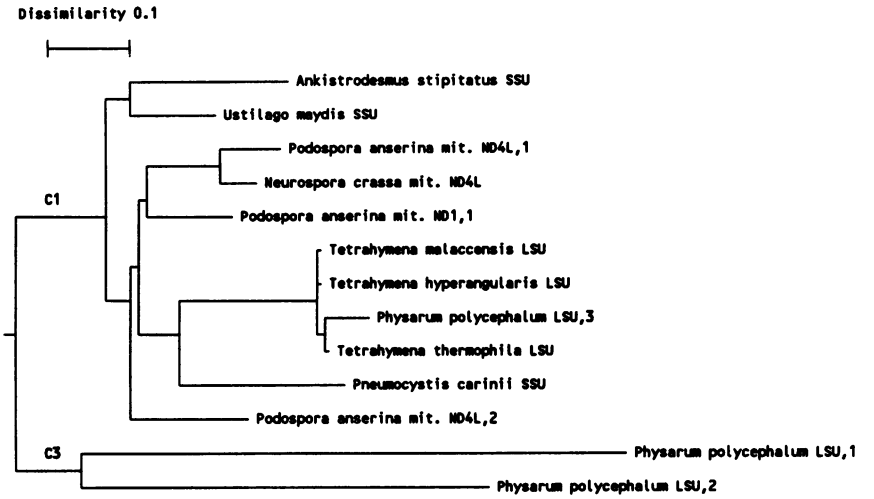

Fig. 3. Relationships among group I introns in nuclear rRNA genes. A tree was constructed on the basis of 83 positions (4) of an alignment of 101 group I intron sequences, but only the cluster comprising the introns situated in nuclear rRNA genes is shown. It includes all the introns classified (4) as C1 plus two of the introns classified as C3, viz. *P. polycephalum* LSU 1 and 2. The scale at the top allows to measure dissimilarity between two sequences (computed according to reference 17) by adding the lengths of connecting branches measured along the horizontal axis. Since the sequences of *T. cosmopolitanis*, *T. pigmentosa*, and *T. sonneborni* LSU rRNA introns are identical to the sequence of the *T. hyperangularis* LSU rRNA intron over the 83 positions taken into account for tree construction, only the *T. hyperangularis* LSU rRNA intron is mentioned. Intron names are abbreviated as in reference 6.

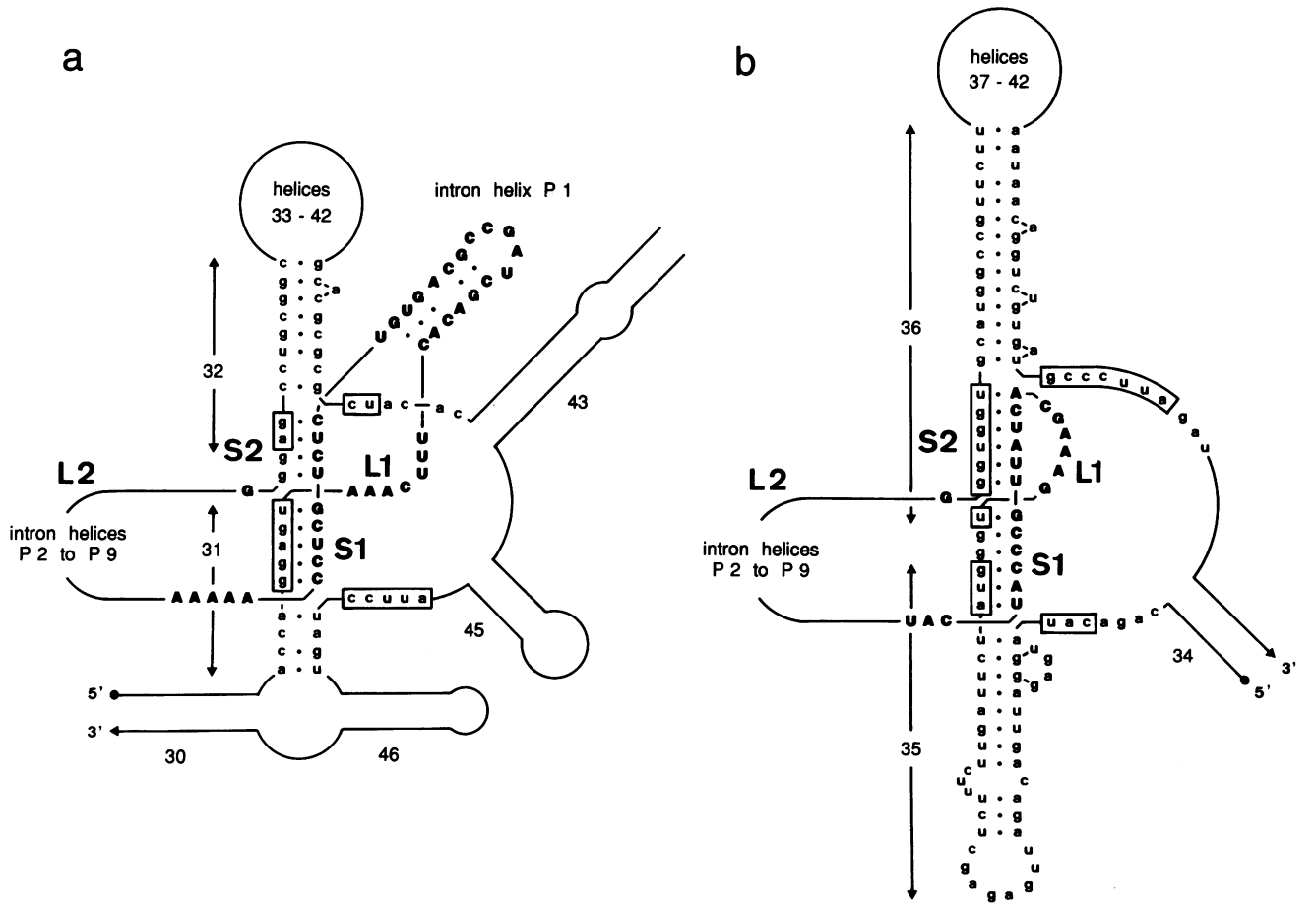

Fig. 4. Model for the secondary structure at the exon-intron boundaries of SSU rRNA precursors. a) Model for *U. maydis*. b) Model for *A. stipitatus*. Nucleotides written in lower case belong to the mature SSU rRNA, those written in bold upper case belong to the intron. SSU rRNA helix numbers correspond to those of Fig. 1, intron helix numbers, preceded by P, correspond to those of Fig. 2a. Stems and loops of the pseudoknot structure formed by the intron and the adjacent sequences are numbered S1, S2, L1, and L2 according to conventions used in reference 28. Boxed sequences are about to pair and to complete helices 31 and 32 (drawing a), or helices 35 and 36 (drawing b) after intron elimination.

with groups discerned by Michel and Westhof (4), viz. B3, C1, C2, and D. Introns belonging to groups A1 and A2, though not entirely resolved, together also form a cluster in our tree. The introns of the remaining groups A3, B1, B2, B4, and C3 are each scattered over a number of different clusters in our tree, in other words, these groups behave as polyphyletic entities in our analysis. On the contrary, all the group I introns hitherto found in nuclear rRNA genes are united in a cluster that comprises the complete subgroup C1, as well as two introns of subgroup C3, as defined by Michel and Westhof (4). The structure of this cluster, which also contains 4 introns in mitochondrial genes for NADH dehydrogenase subunits, is represented in Fig. 3. It shows that the intron most similar to that of *U. maydis* SSU rDNA is the one present in the nuclear SSU rDNA of the green alga *A. stipitatus*. An analysis based on a somewhat larger fraction of the aligned nucleotides confirmed this relationship, but gave slightly different branching patterns for the remaining introns of the cluster. It should be noted that there is a rather large uncertainty on the computation of evolutionary distances if these are based on alignments of limited length, regardless of the method used to process the distance matrix. Nevertheless, the branching pattern of Fig. 3 shows that among the group I introns in nuclear rRNA genes, those of *A. stipitatus* and *U. maydis* SSU are rather distant from the *Tetrahymena thermophila* LSU intron, a finding which conflicts with the view of Dávila-Aponte et al. (12), who consider the *T. thermophila* intron as the one closest related to the intron of *A. stipitatus*.

Structure of the 18S rRNA precursor at the intron boundaries

Since there exist well-documented secondary structure models for SSU rRNA (9) and LSU rRNA (10), it is possible to look for common features in the secondary structure at the intron insertion sites. The 23 group I introns discovered in nuclear, plastid, and mitochondrial LSU rDNA are distributed over 12 insertion sites (6, 8, 27). The four group I introns hitherto found in nuclear and plastidial SSU rDNA are inserted at 4 different sites indicated on Fig. 1. The only feature that these sites seem to have in common is that all of them are situated in relatively conserved areas of primary and secondary structure in both molecules. Three out of 4 insertion sites in SSU rRNA, and 7 out of 12 in LSU rRNA, are located within a helix of the rRNA secondary structure. In *U. maydis* SSU rRNA, the insertion site is located between helices 31 and 32, which are separated by two nucleotides (Fig. 1). At first sight these locations seem rather unfavourable. Indeed, the secondary and tertiary structure of RNA molecules is thought to come about during transcription, but the presence of a large insert with specific structural requirements might result in a distortion of the precursor structure with respect to that found in the mature product, necessitating a rearrangement of this structure after elimination of the intron. Such a rearrangement, which conceivably might be dispensable if the intron is located in a sufficiently large loop of the rRNA secondary structure, seems inevitable if it interrupts a helix strand or is located in an area subject to other structural constraints.

Whereas detailed models have been proposed for the secondary and tertiary structure of the essential components of group I introns (2, 4), we are not aware of a model showing how such an intron fits into the secondary structure of the precursor. Fig. 4a presents just such a model for the relevant area of *U. maydis* SSU rRNA precursor. It shows that helices 31 and 32, which surround the intron insertion site, could actually form a coaxial stack comprising the internal guide sequence of the intron, which brings the splice sites in immediate proximity. The structure forms a pseudoknot where loop L1 (nomenclature according to reference 28) comprises the apex of intron hairpin P1, and loop L2 comprises the rest of the intron structure. In this pseudoknot, some of the base pairs of helices 31 and 32, viz. those adjacent to the multibranched loop that separates them in the final structure (Fig. 1) are disrupted in favour of base pairs formed with the internal guide sequence. The pseudoknot helices S1 and S2 are formed by a hybrid consisting of the ultimate nucleotides of the 5′-strand of helix 31, the internal guide sequence of the intron, and the first nucleotides of the 5′-strand of helix 32. Helices S1 and S2 are extended beyond the pseudoknot by the undisrupted parts of helices 31 and 32. Other examples of pseudoknot helices extended coaxially beyond the pseudoknot loops are known, e.g. in the tRNA-like structure in TYMV RNA (29). The model shows that after elimination of the intron, the bases of helices 31 and 32 previously engaged in pairing with the internal guide sequence only have to form alternative pairs with the remaining bases of their 3′-strands. The structural change is limited to a reduction in size of the multibranched loop connecting helices 31, 32, 43, and 45.

A similar model applicable to the *A. stipitatus* 18S rRNA precursor is shown in Fig. 4b. In this case, the intron insertion site (arrow 2 in Fig. 1) is located in the 5′-strand of helix 36, close to the multibranched loop connecting it to helices 34 and 35. After elimination of the intron, helices 35 and 36, which form a coaxial stack, are completed at the expense of the size of the multibranched loop. Models similar to those of Fig. 4 can be devised for other group I introns and their splice sites in rRNAs, showing that the presence of introns in the precursors, and their elimination by splicing, potentially occurs with minimal rearrangement of secondary and tertiary structure of rRNAs.

ACKNOWLEDGEMENTS

We thank Drs. G.L Hennebert and J.-F. Berny of the Laboratory of Systematic and Applied Mycology of the University of Louvain for supplying and cultivating the *U. maydis* strain. Our research was supported by the Incentive Programme for Fundamental Research in the Life Sciences and the Programme on Interuniversity Poles of Attraction of the Office for Science Policy Programming of the Belgian State, and by the Fund for Medical Scientific Research.

REFERENCES

1. Davies, R.W., Waring, R.B., Ray, J.A., Brown, T.A., Scazzocchio, C. (1982) *Nature* **300**, 719–724.
2. Cech, T.R. (1988) *Gene* **73**, 259–271.
3. Michel, F., Hanna, M., Green, R., Bartel, D.P., Szostak, J.W. (1989) *Nature* **342**, 391–395.
4. Michel, F., Westhof, E. (1990) *J. Mol. Biol.* **216**, 585–610.
5. Burke, J.M. (1988) *Gene* **73**, 273–294.
6. Dujon, B. (1989) *Gene* **82**, 91–114.
7. Woodson, S.A., Cech, T.R. (1989) *Cell* **57**, 335–345.
8. Turmel, M., Boulanger, J., Schnare, M.N., Gray, M.W., Lemieux, C. (1991) *J. Mol. Biol.* **218**, 293–311.
9. Neefs, J.-M., Van de Peer, Y., De Rijk, P., Goris, A., De Wachter, R. (1991) *Nucl. Acids Res.* **19**, 1987–2015.
10. Gutell, R.R., Schnare, M.N., Gray, M.W. (1990) *Nucl. Acids Res.* **18**, 2319–2330.
11. Sogin, M.L., Edman, J.C. (1989) *Nucl. Acids Res.* **17**, 5349–5359.
12. Dávila-Aponte, J.A., Huss, V.A.R., Sogin, M.L., Cech, T.R. (1991) *Nucl. Acids Res.* **19**, 4429–4436.

13. Durocher, V., Gauthier, A., Bellemare, G., Lemieux C. (1989) *Curr. Genet.* **15**, 277–282.
14. Sambrook, J., Fritsch, E.F., Maniatis, T. (1989). *Molecular Cloning. A Laboratory Manual.* 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
15. Hendriks, L., Goris, A., Van de Peer, Y., Neefs, J.-M., Vancanneyt, M., Kersters, K., Hennebert, G.L., De Wachter, R. (1991) *J. Gen. Microbiol.* **137**, 1223–1230.
16. Jukes, T.H., Cantor, C.R. (1969) In Mammalian Protein Metabolism, vol. **3**, pp. 21–132. Edited by H.N. Munro, New York: Academic Press.
17. Van de Peer, Y., Neefs, J.-M., De Wachter, R. (1990) *J. Mol. Evol.* **30**, 463–476.
18. Saitou, N., Nei, M. (1987) *Mol. Biol. Evol.* **4**, 406–425.
19. Barns, S.M., Lane, D.J., Sogin, M.L., Bibeau, C., Weisburg, W.G. (1991) *J. Bacteriol.* **173**, 2250–2255.
20. Van de Peer, Y., Hendriks, L., Goris, A., Neefs, J.-M., Vancanneyt, M., Kersters, K., Berny, J., Hennebert, G.L., De Wachter, R. (in press) *Syst. Appl. Microbiol.*
21. Burke, J.M., Belfort, M., Cech, T.R., Davies, R.W., Schweyen, R.J., Shub, D.A., Szostak, J.W., Tabak, H.F. (1987) *Nucl. Acids Res.* **15**, 7217–7221.
22. Gutell, R.R., Weiser, B., Woese, C.R., Noller, H.F. (1985) *Prog. Nucl. Acid Res. Mol. Biol.* **32**, 155–216.
23. Neefs, J.-M., De Wachter, R. (1990) *Nucl. Acids Res.* **18**, 5695–5704.
24. Muscarella, D.E., Vogt, V.M. (1989) *Cell* **56**, 443–454.
25. Sogin, M.L., Ingold, A., Karlok, M., Nielsen, H., Engberg, J. (1986) *EMBO J.* **5**, 3625–3630.
26. Michel, F., Jacquier, A., Dujon, B. (1982) *Biochimie* **64**, 867–881.
27. Kück, U., Godehardt, I., Schmidt, U. (1990) *Nucl. Acids Res.* **18**, 2691–2697.
28. Pleij, C.W.A., Rietveld, K., Bosch, L. (1985) *Nucl. Acids Res.* **13**, 1717–1731.
29. Pleij, C.W.A. (1990) *Trends Bioch. Sci.* **15**, 143–147.
30. Xu, M.-Q., Kathe, S.D., Goodrich-Blair, H., Nierzwicki-Bauer, S.A., Shub, D.A. (1990) *Science* **250**, 1566–1570.
31. Kuhsel, M.G., Strickland, R., Palmer, J.D. (1990) *Science* **250**, 1570–1573.
32. Nielsen, H., Engberg, J. (1985) *Nucl. Acids Res.* **13**, 7445–7455.
33. Erickson, J.M., Rahire, M., Rochaix, J.D. (1984) *EMBO J.* **3**, 2753–2762.
34. Trinkl, H., Wolf, K. (1986) *Gene* **45**, 289–297.